VERTICAL NAVIGATION DISPLAYS: PILOT PERFORMANCE AND WORKLOAD DURING SIMULATED CONSTANT-ANGLE-OF-DESCENT GPS APPROACHES

Charles M. Oman[2], Andrew J. Kendra[1], Miwa Hayashi[2],
Mary J. Stearns[1], and Judith Burki-Cohen[1]

[1]DOT/Volpe National Transportation Systems Center DTS-79
[2]Man Vehicle Laboratory, Massachusetts Institute of Technology
Cambridge, Massachusetts

ABSTRACT

This study compared the effect of alternative graphic or numeric cockpit display formats on the tactical aspects of vertical navigation (VNAV). Display formats included: a) a moving map with altitude range arc, b) the same format, supplemented with a push-to-see profile view, including a vector flight path predictor, c) an equivalent numeric format, d) a numeric non-VNAV format. Sixteen pilots each flew four different approaches with each format in a Frasca 242 simulator. Vertical and horizontal flight technical errors (FTE), workload, and subjective display preferences were measured. VNAV displays improved vertical FTE by as much as a factor of two without increasing workload. Relative advantages of the graphics formats are discussed.

INTRODUCTION

Aircraft vertical navigation (VNAV) systems provide vertical guidance from or between specified waypoints (AIR-120, 1988; RTCA-SC-159, 1996). Most flight management system (FMS) equipped aircraft have VNAV capability. VNAV displays allow the pilot to plan and check a VNAV route, manually fly the computed path, or monitor VNAV function when the autopilot is flying the aircraft. A recent FAA Human Factors Team report (Abbott, et al., 1996) recommended reduction or eventual elimination of instrument approaches which lack vertical path guidance, in the interests of flight safety, and additional research on VNAV displays. However, the basic human factors requirements for VNAV displays have received little attention.

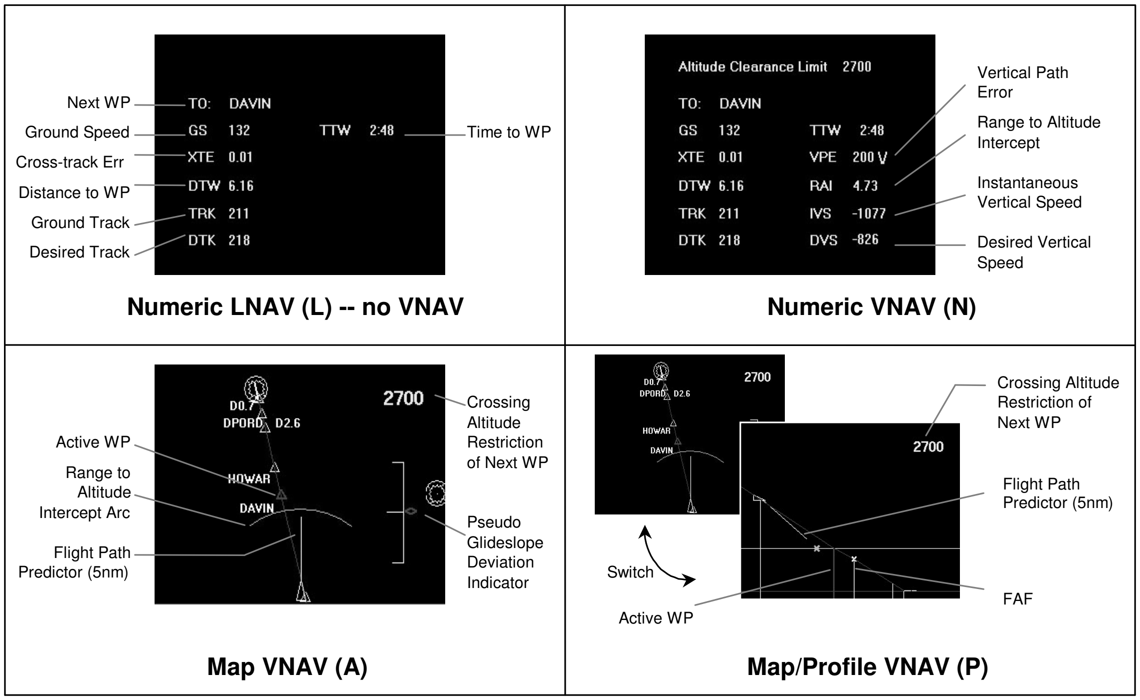
The basic component of existing VNAV displays is a vertical course deviation indicator (CDI), analogous to an ILS glideslope needle. Most systems also display some form of flight path prediction information. The simplest method is to display recommended vertical speed in numeric form. Another is to incorporate vertical prediction features into the graphics "moving map" displays (plan view). For example, on the Boeing/Honeywell FMS displays, a green "altitude range arc" continuously shows where the aircraft will reach a preselected altitude.

Moving map displays, however, do not depict the aircraft's situation relative to terrain or the planned vertical route as explicitly as a profile presentation, which shows altitude vs. distance en route, and cannot be easily used for planning or checking a vertical route. Loss of altitude and terrain awareness led to many fatal controlled-flight-into-terrain accidents (Khatwa & Roelen, 1996). Several manufacturers have already experimented with supplemental "profile" displays. Typically the profile view window is located in a narrow area beneath the moving map window, and it' s limited altitude resolution has been a concern. This may be resolved by overlaying the profile view window on the moving map window, and the pilots switch back and forth between these two windows. The profile display also can include predictor information such as a flight path vector (Gulfstream/ Honeywell G4/5; Hughes, 1995).

The goal of this flight simulator experiment was to evaluate three VNAV display formats representative of the VNAV formats described above (i.e., numeric, map view, and map view with supplemental profile view), and to quantify relative improvement in FTE, workload, and subjective preferences as compared with a traditional display containing only numeric lateral navigation (LNAV, i.e., no VNAV) information.

METHOD

The four display formats evaluated, including one LNAV format, are shown in Fig. 1 below, and were rendered by custom software on an Avidyne 5RR Multifunction Flight Computer. For the three VNAV formats, vertical path deviation was concurrently



displayed on a CDI in the pilot's primary field of view. With the LNAV format, only cross-track error was displayed.

Numeric LNAV format ("L"): Numeric LNAV data, as it might generically appear on a GPS navigation used for non-precision approaches. Pilots had to manage their descent using rule-of-thumb techniques based on the altimeter and vertical speed indicator.

Numeric VNAV ("N"): Similar to the L format, but with additional numeric VNAV data such as vertical path error, necessary to remain on the programmed flight path, and predictor information in the form of recommended vertical speed.

Map VNAV ("A"): A track up moving map navigation display with a green altitude range arc and a magenta line showing the programmed approach route. A 5 mile long velocity vector in front of a fixed triangular aircraft symbol helped pilots judge track angle error, and infer the display scale. Waypoints were shown using a simplified composite of Boeing and SAE symbology. A magenta-colored diamond-shaped "football" repeated the CDI vertical path error information. Map scale was 15 mi. (vertical) × 20 mi. (horizontal) before the final approach fix (FAF), and zoomed by a factor of 3 after the FAF.

Map/Profile VNAV ("P"): A layered display consisting of the "A" format moving map view, replaceable with a profile-view window by pushing a yoke thumb switch. The map view was the default display prior to the FAF, and the profile view became the default after the FAF. The predicted flight path was depicted with a white vector of fixed length. The minimum altitude on the current leg was shown by a yellow horizontal line. Its intersection with the aircraft velocity vector was identified by a moving green cross - corresponding to the altitude range arc. The vertical range of the profile view was 9000 ft. prior to the FAF, and 3000 ft. after FAF. The horizontal scale of the profile view was always equal to the vertical scale of the map view.

Sixteen (960-18000 hr) multi-engine, instrument rated pilots each flew 5-7 practice approaches and 16 trial approaches (4 approach types × 4 display formats) in a Frasca 242 Piper Aztec simulator with patchy turbulence and altitude dependent wind. A computer generated cloud and runway scene was visible out the front window. Four different approach types (JFK R22L, LAX R25L, BOS R33L, and ATH R15L) were used. Each approach type included an initial level flight segment and up to two turns. The order of approach types and display formats was randomized and

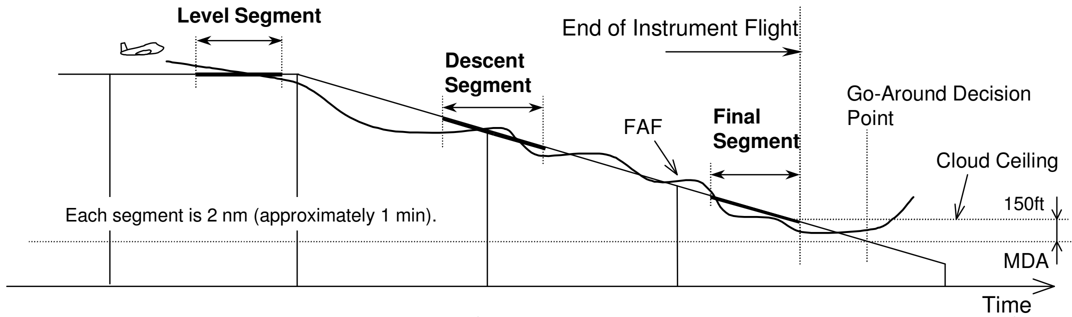


Figure 2: Approach schematic (side view of three 2-mile-long segments - level, descent, and final).

balanced. Pilots were instructed to fly at 120 knots, maintaining a constant angle of descent on all descending legs. As the aircraft approached the minimum descent altitude, pilots made a land or go-around decision. ATC communications were simulated.

After each approach, pilots retrospectively rated their subjective workload before and after FAF using a modified Bedford scale, a measure of spare attention (Roscoe & Ellis, 1990; Huntley, 1993). At the end of the session, pilots completed a three part questionnaire: The first part asked the subjects' opinions of the displays. The second part required the subjects to compare each of the displays on a "head-to-head" (HTH) basis by marking a visual analog "strength-of-preference" scale. The third part asked pilots to rank displays in terms of ease of interpretation (EI), their ability to fly accurately (FA), effect on overall workload (OW), and their overall preference (OP).

RESULTS

Root Mean Square Altitude Error

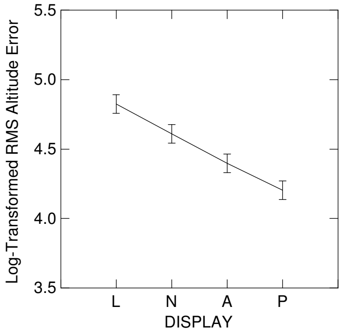
Root Mean Square (RMS) altitude error was computed for each approach. We extracted three two-mile segments from each approach path in order to compare FTE across different approach types. We chose 1) a level segment, 2) a descent segment, prior to the FAF, and 3) a final approach segment, ending at breakout or go-around (see Fig. 2). RMS altitude data were natural-log-transformed to improve normality and analyzed by univariate ANOVA (Systat v. 8.0, SPSS Inc.). Independent variables were subject, approach type, and display format. Subjects and approach types were treated as random effects and display format as a fixed effect. Quasi F-ratios (F' Ratios) were used to approximate the F-ratios for display effect (Winer 1974, Chapter 5).

As expected, a significant effect of subject was found for all three segments (Table 1). The approach effect was also significant except on the level segment, presumably because the level segments were similar across approach types. The display effect, our main interest, also showed significant differences except for the level segment, where the VNAV features were not important. A significant interaction effect of subject and approach was found on the level segment, and of approach and display on the descent segment.

Table 1: Results of Hypothesis Tests on RMS Altitude Errors

Squared multiple (R^2), F-ratio (F), quasi F-ratio (F'), and corresponding p value (p). Only effects with $p < 0.05$ are shown.

Effects	Level Segment ($R^2 = 0.706$)	Descent Segment ($R^2 = 0.654$)	Final Segment ($R^2 = 0.633$)
Subject	$F(15,45) = 10.9$ $p < 0.001$	$F(15,45) = 5.37$ $p < 0.001$	$F(15,45) = 5.76$ $p < 0.001$
Approach	--- ---	$F(3,45) = 8.15$ $p < 0.001$	$F(3,45) = 5.37$ $p = 0.003$
Display	--- ---	$F'(3,9) = 5.65$ $p = 0.019$	$F'(3,8) = 17.4$ $p = 0.001$
Subject * Approach	$F(45,135) = 2.02$ $p = 0.001$	--- ---	--- ---
Approach * Display	--- ---	$F(9,135) = 2.84$ $p = 0.004$	--- ---



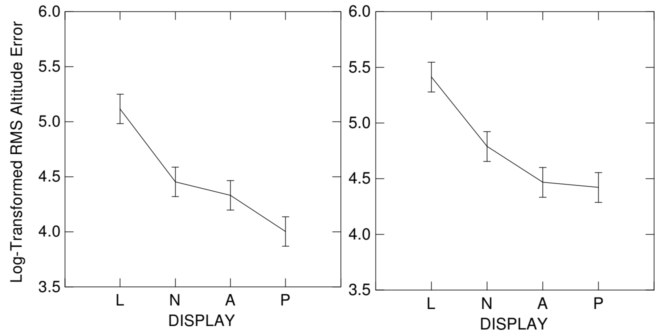


Figure 4, Left: Display effects on least-square means of log-transformed RMS altitude errors <u>on the descent segment</u> within <u>LAX approach type</u>, with ±1 standard error. Right: Same, but within <u>ATH approach type</u>.

Fig. 3 shows the least-square estimates of the mean effect on final segment RMS altitude error of each display type. VNAV information significantly reduced final approach RMS altitude error for all subjects and approaches. Tukey pairwise testing of final segment display effects showed that the graphics VNAV formats (A and P formats) significantly reduced the altitude error as compared to the L format ($p < 0.001$ for A vs. L, and P vs. L). In addition, the P format significantly reduced the altitude error comparing to the N format ($p < 0.001$). Note that the axes in Fig. 3, as well as in Fig. 4, are in log scale, e.g., the least-square mean L format RMS error is $\exp(4.8) \approx 122$ ft and that of the P format is $\exp(4.2) \approx 67$ ft. Thus the P format cut altitude error almost in half.

Due to the significant approach and display interaction effect on the descent segment, Tukey pairwise comparison tests on the display effect were made for each approach separately. Display effects were significant for the LAX and ATH approaches (see Fig. 4), which had the steepest final segment and highest workload demand (see Workload, below), respectively. With the LAX approach, the altitude error was reduced significantly when any one of the VNAV formats was used comparing to L format ($p < 0.040$ for N vs. L, $p < 0.004$ for A vs. L, and $p < 0.001$ for P vs. L). With the ATH approach, the altitude error was reduced significantly when either one of the graphics VNAV formats was used comparing to L format ($p < 0.001$ for A vs. L, and $p < 0.001$ for P vs. L).

Root Mean Square Cross-Track Error

To measure the effect on concurrent lateral FTE, analogous hypothesis tests were performed on the log-transformed RMS cross-track errors. The analysis showed significant main effects of subject and approach type on all three segments[1], but a display effect was found only for the descent segment. The graphical VNAV formats (A and P) reduced the cross-track error significantly when compared to the L format (pairwise Tukey test, $p < 0.001$ for A vs. L and $p < 0.002$ for P vs. L). The A format provided significant improvement in the cross-track error even when compared to the N format ($p < 0.015$ for A vs. N). We suspect this is because the pilot can directly visualize track angle error when using the graphical A or P formats.

On the final segment, a significant interaction effect of approach and display was found. However, pairwise differences between display effects were no significantly different by approach, except on the BOS approach, where the cross-track performance with the A format was better than the N format ($p < 0.017$).

---

[1]. RMS Cross-Track Errors -- Level segment: Subject $F(15, 45) = 5.45$, $p < 0.001$; Approach $F(3,45) = 24.1$, $p < 0.001$; Subject * Approach $F(45,135) = 1.49$, $p < 0.043$; Display, trial number, and all other interaction terms not significant.
Descent segment: Subject $F(15, 45) = 4.47$, $p < 0.001$; Approach $F(3,45) = 40.2$, $p < 0.001$; Display $F'(3,10) = 7.39$, $p < 0.007$; Subject * Approach $F(45,135) = 1.84$, $p < 0.004$; trial number, and all other interaction terms not significant.
Final segment: Subject $F(15, 45) = 7.80$, $p < 0.001$; Approach $F(3,45) = 49.8$, $p < 0.001$; Subject * Approach $F(45,135) = 1.65$, $p < 0.015$; Approach * Display $F(9,135) = 2.40$, $p < 0.015$; Display, trial number, and all other interaction terms not significant.

Workload

Average workload scores for all displays before and after the FAF fell in the middle of the "tolerable" range on the Bedford Scale. ANOVA and the hypothesis testing showed significant effects of subject and approach trial number both before and after the FAF, and an approach effect only before the FAF[2]. A Tukey pairwise comparison test on the approach effect before the FAF showed that the ATH approach significantly increased the pilot' s workload score compared to other approach types[3]. No significant display effect was found either before or after the FAF, suggesting that the performance improvements associated with VNAV displays as compared to traditional LNAV/altimeter did not come at the expense of any increase in workload.

Subjective Preferences

To analyze the post-session questionnaire data, HTH preference indications were converted to a numeric score using a tournament scoring method, and then converted to ranks. Rank sums across all 16 pilots for each of the 5 preference measures were computed. In terms of overall preference, subjects ranked the P format as best in both tournament (HTH) and direct (OP) measures. The A format ranked first in terms of ease of interpretation (EI) and effect in reducing workload (OW). Subjects believed that they could fly most accurately using the P format (FA), and that was consistent with their FTE data. Friedman ANOVA showed the 5 measures were significantly concordant (Friedman test statistic = 13.6, df = 3, $p < 0.004$).

All pilots said that the map view should be the default prior to the FAF, but 9 of 16 responded that the map should remain as the default after passing the FAF, and the profile view should remain "push-to-see." 13 of 16 agreed that the map/profile switch should be mounted on the yoke rather than on the instrument panel. 11 of 13 pilots said they never or rarely had any problems interrelating the map and profile views. Several pilots suggested that waypoint names and numeric altitudes be shown in both map and profile views.

The N format was ranked third on all 5 preference scales. 10 of 16 pilots reported they flew the N format by comparing the Instantaneous Vertical Speed (IVS) and Desired Vertical Speed (DVS) values rather than distance measures such as Distance to Waypoint and Range to Altitude Intercept. Several commented they found the IVS/DVS presentation more intuitive, and noted they could cross check IVS with their barometric instrument.

Most pilots preferred the graphic formats (A and P) over the numeric formats (L and N), but some pilots strongly preferred the numeric formats.

CONCLUSIONS

The analysis of RMS altitude error data showed that the three VNAV formats we tested significantly reduced vertical FTE on final segment of simulated constant-angle-of-descent approaches as compared to traditional methods employing LNAV data (L format), an altimeter, and rules-of-thumb. The improvement in vertical FTE on final segments associated with the Map/Profile format (P format) approached a factor of two. Graphical formats were particularly helpful on the descent segments in two relatively challenging approaches of the four flown. On final approach, performance with the Map/Profile format was significantly better than the numeric format (N format).

Displays also affected lateral FTE. Cross-track error was significantly reduced when using the graphic formats during the descent segments. This was expected, since these formats provided a moving map (plan) view. With the Map/Profile format, the profile view became the default after the FAF since horizontal maneuvers are not normally made on final approach, but the majority of our pilots suggested to keep the map view as the default throughout the entire approach, with the profile view remaining "push-to-see."

Our results provided quantitative support for the widely held view that map with altitude range arc allows vertical performance equivalent to profile displays. However, supplementary profile views probably help improve vertical situation awareness. We demonstrated that a profile view can be layered in the same display space as a map view, usually without any increase in lateral or vertical FTE as compared to the map view alone.

---

[2]. Workload before FAF: Subject $F(15,45) = 9.05$, $p < 0.001$; Approach $F(3,45) = 7.04$, $p < 0.001$; trial number $F(1, 134) = 24.1$, $p < 0.001$; Display and all interaction terms not significant.
Workload after FAF: Subject $F(15,45) = 4.81$, $p < 0.001$; trial number $F(1, 134) = 12.8$, $p < 0.001$; Approach, Display and all interaction terms not significant.

[3]. $p < 0.018$ for ATH vs. JFK, $p < 0.001$ for ATH vs. LAX, and $p < 0.046$ for ATH vs. BOS

No significant difference was found in workload scores across display formats. Our pilots were anxious to do well, and they may have worked about as hard as they could in each case to keep the workload constant and allow the performance vary. This may explain why their workload scores appeared constant across displays.

ACKNOWLEDGMENTS

Supported by FAA Office of the Chief Scientific and Technical Advisor for Human Factors, AAR-100 via DOT/Volpe National Transportation Systems Center, Cockpit Human Factors Program, DTS-79 Contracts DTRS-57-P-80766 and DTRS-57-97-C-00022-TTD02. We thank our 16 subjects, and also M. Pettitt and T. Imrich of FAA, D. Sussman, R. Shih, M. Bud, F. Sheelen, J. Giurleo, and J. Turner of Volpe, and A. Natapoff, A. Rasmussen, C. Carr, and N. Patrick of MIT.

REFERENCES

Abbott, K, Slotte, S, & Stimson, D (1996). The Interfaces Between Flightcrews and Modern Flight Deck Systems (FAA Human Factors Team Report): FAA.

FAA AIR-120. (1988). AC 20-129 Airworthiness Approval of Vertical Navigation (VNAV) systems for use in the US National Airspace System (NAS) and Alaska (Advisory Circular): FAA. US Department of Transportation, Publications Section, Washington, DC.

Hughes, D (1995a, February 6, 1995). Gulfstream using vertical profile display. Aviation Week and Space Technology, 142, 55.

Huntley, MS Jr. (1993). Flight Technical Error for Category B Non-Precision Approaches and Missed Approaches Using Non-Differential GPS for Course Guidance (Report DOT-VNTSC-FAA-93-17). Cambridge, MA: US DOT Volpe National Transportation Systems Center.

Khatwa, R., & Roelen, A (1996). An analysis of controlled-flight-into-terrain (CFIT) accidents of commercial operators, 1988 through 1994. Flight Safety Digest, 15(4/5), 1-45.

Roscoe, A, & Ellis, G (1990). A subjective rating scale for assessing pilot workload in flight (Tech. Report TR90019). Farnborough, U.K.: RAE.

RTCA-SC-159. (1996). Minimum operational performance standards for global positioning system/wide area augmentation system airborne equipment (MOPS RTCA/DO-229). Washington, DC: RTCA, 1140 Connecticut Ave, NW, Suite 1020 Washington, DC 20036.

Winer, B (1971). Statistical Principles in Experimental Design McGraw-Hill Book Company, New York, NY.